# INVITED

**#01**

**Invited expert meeting on revising U.S. EPA's guidelines for deriving aquatic life criteria, September 14-16, 2015, Arlington, VA**

**Name of presenter, organization, address, phone, email:**
Tom Augspurger, U.S. Fish and Wildlife Service, [redacted]

**Presenter biography (300 word max):**
[redacted]

**Presentation title:**
Derivation and application of taxon-specific criteria: Additional resolution in WQC recommendations

Tom Augspurger (USFWS, [redacted]), Chris A Mebane (USGS, [redacted] Ning Wang, John M Besser, Chris G Ingersoll (USGS, [redacted]), and Sandy Raimondo (USEPA, [redacted]

**Abstract (500 words max):**
In 2005, a conceptual approach for deriving taxon-specific criteria was presented to USEPA's Science Advisory Board Aquatic Life Criteria Guidelines Consultative Panel. Conceptually, a taxon-specific criterion could be derived to protect a species, genus, or family that is not

adequately protected by general national aquatic life ambient water quality criteria (WQC). Taxon-specific criteria may be appropriate for taxa at the lower tail of species sensitivity distributions or to focus on taxa of special management concern (e.g., species of recreational, commercial, ecological or cultural importance). The benefit of including taxon-specific criteria in WQC is the provision of technically sound risk management options for water quality managers' consideration when developing state or tribal water quality standards. Taxon-specific criteria would complement the general national WQC which are derived to be protective of a large number of taxa but not meant to protect all species.

We illustrate the technical feasibility and management applications of a taxon-specific criterion for freshwater bivalve mollusks and ammonia. In USEPA's 2013 WQC for ammonia, the 10 lowest (most sensitive) species mean acute values (SMAVs) and two lowest species mean chronic values (SMCVs) are for mussels. It is therefore reasonable to expect some of the approximately 270 species of mussels and 35 species of clams in the USA not included in the dataset may be equally or more sensitive. Further, widespread distribution and protected status of many mussel species indicate taxon-specific estimates of safe concentrations may be valuable to natural resource managers. We added subsequently published toxicity data for bivalves to the 2013 ammonia WQC database. Data synthesis followed the USEPA 1985 guidelines for deriving WQC, with exception that only data for bivalves were used and that SMAVs and SMCVs were retained rather than collapsed into genus means. From 22 SMAVs, a bivalve-specific final acute value (FAV) of 22.6 mg/L total ammonia nitrogen (TAN) at pH 7 and 20$^{o}$C was derived. The FAV was divided by 2 (after examination of dose-response slopes ensured this was appropriate) for a bivalve-specific maximum instream concentration of 11.3 mg TAN/L at pH 7 and 20$^{o}$C. From four SMCVs for bivalves, a bivalve-specific acute-chronic ratio of 21.6 was applied to the FAV yielding a bivalve-specific instream chronic concentration of 1.0 mg TAN/L at pH 7 and 20$^{o}$C. At the example pH 7 and 20°C, the bivalve-specific acute concentration is 1.5-fold lower than the 2013 WQC acute criterion magnitude and the bivalve-specific chronic concentration is 1.9-fold lower than the 2013 WQC chronic criterion magnitude. Taxon-specific criteria could also tailor duration and frequency recommendations to the taxon of interest.

When problem formulation in deriving national WQC indicates that certain taxa may be sensitive to that particular chemical, of special biodiversity concern, or management significance, deriving taxon-specific criteria can provide risk managers with options focused on protecting those taxa. Inclusion of taxon-specific criteria within WQC recommendations may also facilitate timely completion of Endangered Species Act consultations by providing probabilistic estimates of hazard based on the subset of data most relevant to individual taxa of conservation concern.

**#02**

Augmenting Species Diversity in Water Quality Criteria Derivation using Interspecies Correlation Models

Mace G. Barron, Crystal R. Lilavois, Morgan M. Willming, Jill Awkerman, Sandy Raimondo

The specific requirements for taxa diversity of the U.S. EPA 1985 ambient water quality criteria (AWQC) guidelines have limited the number of criteria developed for aquatic life protection. EPA developed the Web-based Interspecies Correlation Estimation (Web-ICE) tool to allow extrapolation of acute toxicity in commonly tested species to endangered species and other taxa with limited toxicity data. ICE models are log-linear least squares regressions that estimate acute toxicity (LC50/LD50) to an untested species, genus, or family (predicted taxon) from known toxicity of a single surrogate species. Approximately 2400 ICE models have been developed and validated. Models are available for aquatic animals, algae, and terrestrial wildlife using the online Web-ICE tool (http://epa.gov/ceampubl/fchain/webice/), which also includes modules for endangered species extrapolation and generation of species sensitivity distributions (SSDs). Web-ICE toxicity estimates and SSD-based hazard concentrations have been demonstrated to have high accuracy for closely related taxa (within same order).
Published and ongoing research shows that ICE model toxicity extrapolation can be used to increase taxa diversity in SSDs necessary for the development of AWQC with reasonable uncertainty.

**#03**

Michael Bock, Ph.D.
Ramboll Environ
[redacted]

[redacted]

**Statistical Tools to Evaluate Species Sensitivity Distributions and Calculate Final Acute and Chronic Values**

The derivation of the final acute value (FAV) from toxicity data is an integral step in the development of aquatic life criteria. The FAV is based on an evaluation of the distribution of the toxicity data, which is a form of species sensitivity distribution (SSD). SSDs are a tool by which multiple toxicity values can be integrated to estimate a concentration predicted to be protective of a given percentage of species. A variety of factors can affect estimates of the FAV based on SSDs, including the size of the dataset, the representativeness of the species in the dataset, the quality of individual toxicity data points, and the statistical methods used to calculate the FAV. Current guidance defines the FAV as the 5$^{th}$ percentile of the SSD calculated using the method developed by Erickson and Stephan (1998). The 1998 method assumes that the toxicity data follow a log-triangular distribution. It uses a simplified calculation to derive the 5$^{th}$ percentile, which is ordinarily based on the four lowest toxicity values from the toxicity distribution. Although the 1998 method was innovative at the time it was issued, advances in computing power and SSD analytical methods (e.g., Aldenberg, 2002; Newman 2000) offer promise for application of more statistically rigorous methods in developing FAVs. Alternative methods that will be discussed in this presentation include (1) the use of regression analysis to fit the fraction of species affected and the toxicity value to a theoretical distribution, (2) the use of the univariate maximum likelihood estimator (MLE) to fit the data to a theoretical distribution which avoids potential issues in determining the fraction affected, and, (3) non-parametric bootstrap methods. These methods do not rely on the simplifying assumption of a log-triangular distribution and allow all of the data to be used to derive percentiles. These methods can be applied to datasets that follow a variety of distributions (normal, log-normal, Weibull, logistic, log-logistic, etc.), including in some cases datasets that do not fit any of the target distributions (non-parametric). The alternative methods to be discussed are readily available and are easily applied. In addition to exploring several alternative methods, we will discuss the pros and cons of regression based methods versus univariate distribution fitting based methods. We will demonstrate how these methods can be applied to the development of FAVs and compare their performance to the Erickson and Stephan (1998) method. Topics to be covered include (1) evaluating and managing data, (2) determining the distribution that best describes the data based on goodness of fit testing and graphical methods, (3) fitting the data to a distribution or regression based model, and (4) using the fitted model to calculate the FAV. We will discuss uncertainty, sample size, and methods to avoid bias in the underlying toxicity data and we will compare these methods to the Erickson and Stephan (1998) method.

#04

**The Use of Multi-Linear Regression to Derive Site-Specific Water Quality Criteria for Metals: A Complementary Approach to the Biotic Ligand Model**

Kevin V. Brix[1], David K. DeForest[2], Lucinda M. Tear[2], Robert C. Santore[3], Adam C. Ryan[3], and William J. Adams[4]

[1]EcoTox, Miami, Florida
[2]Windward Environmental, [redacted]
[3]Windward Environmental, [redacted]
[4]Rio Tinto, [redacted]

Over the past 20 years there has been a concerted effort to develop Biotic Ligand Models (BLMs) for the derivation of water quality criteria (WQC) and site-specific water quality standards for metals and for application in ecological risk assessments. While BLMs have largely been accepted in Europe and to a lesser extent in other regions, implementation of BLMs has been quite limited in North America. In 2007, the USEPA recommended a BLM-based WQC for Cu in freshwater, but few states have adopted this criterion, and typically only under site-specific conditions. As states have considered adopting the BLM-based WQC, the perceived complexity and data requirements of the model create new challenges for how to implement the approach. For example, not all receiving waters in a state may have suitably comprehensive chemical monitoring to provide all the parameters required by the BLM. An implementation plan for BLM-based WQC, therefore, may need to consider a tiered approach with alternative strategies suitable for sites where data limitations prevent use of the BLM or where state regulators are concerned that the BLM is too complex for routine implementation. As one such alternative, we developed a complementary approach to the BLM, which begins by using the BLM framework to identify key water chemistry parameters influencing metal toxicity. These key parameters were then used to develop multi-linear regression (MLR) models that predict metal toxicity on a species-specific basis. Step-wise linear regression based on Akaike's Information Criterion (AIC) was used to identify which water chemistry parameters and interaction terms resulted in the best models for predicting toxicity. Species-specific MLRs were then pooled in a manner analogous to the pooling of slopes used in previous US-based hardness- dependent WQC. The resulting final criteria equations are equivalent in form with hardness- dependent WQC equations, but with several additional terms (typically dissolved organic carbon, pH, hardness, and several interaction terms). Comparison of acute and chronic MLR-based to BLM-based WQC for Al, Cd, Cu, Ni, Pb, and Zn indicates MLR performance is generally comparable to the BLM (BLM-based WQC for Al, Cd, Ni, Pb, and Zn were derived following USEPA's approach for BLM-based Cu WQC). We suggest that the MLR-based approach, which includes the mechanistic foundation of the BLM for selecting relevant water quality parameters and interactions between parameters, but requires fewer parameters and has an output

similar to the widely acceptable hardness-dependent WQC, may be a useful complementary tool in a tiered framework for derived WQC for metals. This approach may also be useful in developing WQC for other contaminants where multiple water quality characteristics influence the bioavailability and/or toxicity of the chemical.

Presenter: Kevin V. Brix, Ph.D.
Organization: EcoTox
[redacted]

[redacted]

#05

**David Buchwalter**

North Carolina State University, Department of Biological Sciences
[redacted]

**Presenter Biography:** [redacted]

**Presentation Title:** Explanations for the disconnection between the laboratory and field in understanding the effects of metals on aquatic insects.

**Abstract:** The distributions of species that populate natural systems are often quite different in composition from the distributions of species that populate toxicity datasets. Particularly noteworthy is the extent to which aquatic insects are under-represented in toxicity datasets relative to their importance in nature and in ecological monitoring programs. Remarkably, data from only 1 insect species is required for the generation of water quality criteria despite the fact that there are over 7,000 known aquatic insect species in North America alone. When aquatic insect data are examined for trace metal toxicity, it would appear that they are generally insensitive to metals. However, aquatic insects are often some of the first species to disappear from metal-contaminated sites. In fact, typical laboratory results would indicate that insects only respond to dissolved metals at concentrations orders of magnitude larger than those found in the most insect-depleted contaminated sites. Even with mounting evidence highlighting the obvious disconnect between laboratory toxicity tests and field observations regarding metal toxicity to aquatic insects, water quality criteria for metals continues to rely primarily on toxicity values derived from short term dissolved-only exposures. In this talk, I will discuss four key reasons why such tests don't provide relevant data for this important faunal group, focusing upon recent advances in our understanding of bioaccumulation and mechanisms of toxicity. I will also discuss our work on phylogenetic extrapolations and the development of a mayfly model for aquatic toxicology.

**Key words:** Dietary toxicity, Ecological representation, Phylogenetic extrapolation, Mayfly model development

**#08**

William H. Clements
Department of Fish, Wildlife, and Conservation Biology
Colorado State University
[redacted]

**Biological Sketch:** [redacted]

**Integrating mesocosm experiments and field data to develop water quality criteria for contaminants in aquatic ecosystems**

**Abstract:** The limitations of single species toxicity tests for predicting effects of contaminants on aquatic ecosystems have been well described in the literature. Standards or criteria derived exclusively from laboratory experiments may not be protective of natural benthic communities and therefore should be supported by alternative approaches. For certain classes of stressors in which direct toxicological effects are unlikely (e.g., nutrients, suspended sediment, deposition of iron colloids) or in which standard test organisms are highly tolerant (total dissolved solids), field-derived exposure-response relationships have been used to establish standards. The major challenge using field-derived data is to separate effects of the specific stressor of interest from other potentially confounding factors. Mesocosm experiments provide an ecologically realistic alternative to laboratory toxicity tests while controlling for the confounding variables associated with field-based approaches. In addition to providing mechanistic insights into stressor-response relationships across different levels of biological organization, mesocosm experiments can be coupled with field assessments to address important policy issues. In this presentation I will review results of two large scale studies in which mesocosm experiments were integrated with field-based approaches to validate current water quality criteria or benchmarks.

Using field data from spatially extensive surveys of over 300 Colorado streams, we established concentration-response relationships between trace metals and macroinvertebrate community structure. Results showed consistent and predictable alterations in community composition along a gradient of metal contamination. These data were supported by a set of 24 stream mesocosm experiments that established concentration-response relationships and allowed us to estimate community-level EC20 values for metals. Additional evidence for a causal relationship between metals and macroinvertebrate responses was provided by a long-term (24 year) "natural" experiment in which we documented macroinvertebrate responses to the removal of metals. Although these data showed significant improvements in water quality and macroinvertebrates over time, communities remained impaired when metal concentrations exceeded the community-level EC20 values.

In the second example mesocosm experiments were used to test the validity of a recent field-based benchmark for conductivity proposed by the U.S. EPA. This 300 µS/cm benchmark has been criticized because of the potential influence of confounding variables and the difficulty demonstrating a causal relationship between elevated conductivity and benthic macroinvertebrate responses. We conducted a series of stream mesocosm experiments to quantify the effects of several major ions on natural aquatic insect assemblages. Results showed considerable variation among endpoints and salts; however, we observed significant conductivity-response relationships for all major ions tested. Although EC20 values were greater than the 300 µS/cm benchmark for most endpoints, several metrics were affected at conductivity levels near or below this benchmark. We also observed significant variation in sensitivity between communities from different streams, suggesting the importance of accounting for context-dependent responses when establishing benchmarks for contaminants. These findings demonstrate the strength of integrating descriptive, field-based approaches with mesocosm experiments to establish causal relationships and to identify safe concentrations of contaminants in aquatic ecosystems.

#09

**Invited Expert Meeting on Revising U.S. EPA's Guidelines for Deriving Aquatic Life Criteria**

**14-16 September 2015, Arlington, VA**

**Name of Presenter**: S. M. Cormier

**Organization:** NCEA[redacted], ORD, EPA

**Address, City, State, Zip:** [redacted]

**Telephone:** [redacted]

**Email Address**: [redacted]

**Presenter Biography** (300 words max)

[redacted]

**Presentation Title:** Field-based Methods for Developing Water Quality Benchmarks

**Abstract** (500 words max)

The theoretical foundation of Aquatic Life Criteria (ALC) is based on the premise that organisms have evolved different physiologies and behaviors. Variation within and among species results in different tolerances for both natural and xenobiotic challenges. Sensitivity distributions of tolerance ranges for an agent characterize taxa that are at risk at different concentrations. The background of xenobiotics is zero (e.g., DDT), therefore, the effect endpoints of nationally representative species are useful for identifying protective values. The background of naturally occurring agents (e.g., salts, metals) is variable; and therefore, the representative species are those that inhabit areas with a similar background concentration.

Based on these fundamentals, field-based methods were developed for ALC and applied to anthropogenically increased concentrations of major ions measured as specific conductivity (SC). Where a tolerance is exceeded, a genus is not expected to be present; therefore, the upper tolerance levels of SC can be used to develop a model that predicts the proportion of taxa extirpated at a given exposure.

Paired measurements of SC and biota within a Level III Ecoregion were used to estimate the exposure level that potentially extirpates each of more than one hundred benthic aquatic invertebrates. The upper tolerance limit for each genus observed $\geq$25 times in the ecoregion was estimated at the 95$^{th}$ centile of observations of a genus, an extirpation concentration (XC95). The XC95 values were used to develop a sensitivity distribution. The SC level that affects 5% of the genera was calculated at the 5$^{th}$ centile of the sensitivity distribution (HC05). The model was validated with an independent data set and met the criteria of probable causation and minimal confounding. This method is an EPA-approved method for developing Water Quality (WQ) benchmarks.

Because a genus will rarely occur where the natural regional background exceeds the upper tolerance limit for that genus, the lower possible tolerance limit of genera in a region is defined by the natural background. Theoretically, there will be a regular relationship between natural background SC and the SC that affect 5% of genera. To test this deduction from theory, the SC levels affecting upper 5% of genera were modeled with the background SC for several Level III Ecoregions in the United States.

Background SC was estimated at the 25$^{th}$ centile of all sampled sites in each ecoregion. A least squares regression of HC05 values against background yielded a strong statistical relationship. The relationship between the salt-intolerant genera and background SC demonstrates that salt-intolerant genera occupy the lowest available SC niche in an ecoregion. The regression model makes it possible to use SC background in an ecoregion to predict the level that is expected to extirpate organisms adapted to low SC niches in an ecoregion, i.e., to predict a WQ benchmark from background. The views expressed in this abstract are those of the authors and do not necessarily represent the views or policies of the U.S. EPA.

**Key Words:** conductivity, maximum tolerance, field-based methods, salt-intolerant niche, sensitivity distribution, streams, benthic macroinvertebrates, ionic concentration, major ions, field-based criteria, assessment of criteria

#10

## Ambient Water Quality Criteria: Protectiveness of Threatened and Endangered (T&E) Species and Aquatic-dependent Wildlife

David K. DeForest*[,1], Robert W. Gensemer[2], Carrie Claytor[2], Steven P. Canton[2], Robert C. Santore[3]

[1]Windward Environmental, [redacted]
[2]GEI Consultants, [redacted]
[3]Windward Environmental, [redacted]

* Presenter: David DeForest
Windward Environmental
[redacted]

**Biography**

[redacted]

**Abstract**

The Environmental Protection Agency's (EPA's) 1985 guidelines for ambient water quality criteria (AWQC) development do not specifically refer to protection of aquatic threatened and endangered (T&E) species, and only limited guidance is provided with regard to ensuring protection of aquatic-dependent wildlife. With respect to T&E species, the 1985 guidelines note that criteria should be protective of commercially, recreationally, and other “important” species. This generally means that if a species mean acute value (SMAV) or species mean chronic value (SMCV) for an "important" species is less than the final acute value (FAV) or final chronic value (FCV) for a given substance, then the FAV and/or FCV may be lowered. However, because the goals of the Clean Water Act differ from those of the Endangered Species Act, such criteria may not necessarily equate to levels of protection expected for T&E species. Regarding aquatic-

dependent wildlife, the 1985 guidelines include calculation of a final residue value (FRV), which is the lowest maximum permissible tissue concentration for wildlife (or human health) divided by a bioconcentration factor (BCF) or bioaccumulation factor (BAF). We propose using a risk-based framework for evaluating whether AWQC are protective of T&E species and aquatic-dependent wildlife. This includes development of (1) a problem formulation for defining the analysis plan; (2) an exposure characterization for describing the receptors and exposure pathways that will be evaluated; (3) an effects characterization for describing and presenting the toxicity data for T&E species, aquatic-dependent wildlife, and surrogates; and (4) a risk characterization in which the protectiveness of the AWQC relative to T&E species and aquatic-dependent wildlife is evaluated. We will use cyanide as an example, as the current cyanide AWQC are based on the sensitivity of rainbow trout (*Oncorhynchus mykiss*), which is considered a commercially and recreationally important species. In addition, cyanide is not a bioaccumulative substance, so it requires a framework that is different from the EPA's current FRV approach. For the T&E evaluation, we propose a weight-of-evidence approach that considers: (1) the availability of empirical cyanide toxicity data for T&E species and intra-genus and intra-family T&E surrogate species; and (2) estimated cyanide toxicity data for T&E species and surrogates using the EPA’s Interspecies Correlation Estimations (ICE) model. In this example, in large part due to the very steep concentration-response curve for cyanide, division of the rainbow trout SMAV by two to calculate the acute criterion results in a very low effect level (i.e., the estimated LC01 is more than one-half the LC50). Overall, this analysis concluded that the freshwater cyanide AWQC are likely to be protective of most T&E species, although, based on the ICE model, some individuals of two T&E species could be affected at acute criteria concentrations. Regarding aquatic-dependent wildlife, predicted cyanide doses at surface water concentrations equal to AWQC were below conservative thresholds for surrogate birds and mammals. Overall, incorporation of such a framework in revised guidelines for AWQC development would help to standardize the tools and methodology for ensuring that criteria are protective of T&E species and aquatic-dependent wildlife.

**#11**

Charles Delos

Great Lakes Environmental Center

[redacted]

[redacted]

[redacted]

[Presenter Biography]

[redacted]

[Title and Abstract]

A Unified Aquatic Life Framework for Addressing the Affected Percentages of Individuals, Species, and Time

The concept of a Species Sensitivity Distribution (SSD) has served aquatic life criteria development since 1980 but has some limitations as currently applied. The existing perspective (a) ignores information on the level of effect in any species other than the hypothetical 5$^{th}$ percentile, and (b) offers little rigor for addressing time variability: there is no Guidelines procedure for deriving the target attainment frequency. With the existing Guidelines, time variability of exposure is not addressed during the derivation of criterion concentrations, but is an appendage added after the fact.

This proposal describes a unified tiered approach for quantifying risks of effects on aquatic life, simultaneously addressing both the concentration and time components of criteria. For each species in an SSD, an "Assemblage Toxicity Index" (Erickson 2008) applies the concentration-response curves to quantify the percentage of effect at any particular concentration. Such percentages can be aggregated across species to determine the total effect the particular concentration has on the assemblage. Then, considering realistic degrees of temporal (or

spatial) variability in ambient waters, the distribution of concentrations is integrated over time (or space) to obtain the aggregated effects. This approach requires no data beyond what EPA already has available (but does not necessarily use) for deriving aquatic life criteria. This first and most important tier of the proposed framework thus discerns the overall level of risk for every species in the tested assemblage, and can be applied as well to endangered species having estimated rather than tested sensitivity.

The second proposed tier unifies acute and chronic perspectives using a three-parameter kinetic toxicity model. Such an approach requires no new toxicity information beyond what is already available: an *acute* EC50, a *chronic* EC20, and the slope of the concentration-response curve, such as given by the ratio the *chronic* EC50/EC20. Applying this approach requires a time sequence of exposure concentrations, not merely a probability distribution of concentrations; however, long synthetic time series can easily be generated to mimic real-world time series in their variability and serial correlation. This approach calculates the level of stress (relative to that causing any particular percentage of effects in a species) through time for unsteady exposures that are neither acute nor time-invariant chronic.

The third tier links a life-stage-specific population model to the toxicity model to predict the effect on populations, either expressed as population growth rate (density-independent model) or population density (density-dependent model). This allows discernment of (a) differing sensitivities of different life stages, and (b) differing rates at which affected individuals can be replaced in different species (e.g., k- selected versus r-selected species). For a small selection of representative surrogate taxa, this tier requires species life-history information beyond what has been collected during past criteria derivations.

The above approaches have already been applied in EPA. EPA (2011) applied the first tier to a single bird species (selenium reproductive effects); Erickson (2008, 2012) applied it to a plant assemblage (atrazine growth effects). Delos (2008) applied all three tiers to an aquatic animal assemblage (ammonia survival and reproductive effects).

**#12**

Abstraction Submission for:
"Meeting on Revising U.S.EPA's Guidelines for Deriving Aquatic Life Criteria"

Title:
"Beyond Species Sensitivity Distributions: More Fully Characterizing Toxicity to Assemblages of Aquatic Taxa"

Russell Erickson
U.S. Environmental Protection Agency
Mid-Continent Ecology Division, [redacted]

Assessments of chemical toxicity to aquatic organisms often employ statistical distributions of toxic effect concentrations across an assemblage of taxa of interest, generically referred to as "species sensitivity distributions" (SSDs). SSDs contribute to ecological risk assessments by identifying the fraction of taxa potentially affected at a designated exposure concentration and have been an important and useful part of aquatic risk assessments for over 30 years. However, the information they provide has some limitations. First, each point on an SSD identifies a concentration eliciting a single level of effect for a particular taxon; no information is provided for effects on other taxa at this concentration or on effects on this taxon at other concentrations. Second, each effect concentration constituting an SSD addresses just a single point in time, typically following constant exposure; consideration is not given regarding how effects would differ for other exposure time-series. Third, SSDs are typically subject to a variety of statistical estimation problems, especially regarding the meaning and uncertainty of percentiles below that of the most sensitive taxon, because techniques premised on random sampling are applied to very nonrandom samples and simple distributions are imposed on data with more complicated underlying shapes and multiple sources of uncertainty. Finally, the percentiles of SSDs have uncertain relationships to tangible effects in aquatic communities which the species assemblage is intended to represent; efforts to establish this relationship have had mixed results. To better address these issues, U.S.EPA has given attention to an "assemblage toxicity index" (ATI) that more fully utilizes ecotoxicology data to quantify the impact of a toxic chemical on an assemblage of species, addressing both level of effect and effect of time. This index, like an SSD, is one dimensional and thus still sacrifices certain ecotoxicological information, and also has statistical estimation and uncertainty issues that must be recognized and addressed. However, for some applications, it can provide advantages over SSDs. To illustrate this approach, effects of atrazine on aquatic plants will be discussed. An ATI based on growth in single-species toxicity tests with atrazine will be presented, and the use of experimental ecosystem tests to set a level of concern for this index will be described.

Russell Erickson
Mid-Continent Ecology Division
National Health and Environmental Effects Research Laboratory

U.S. Environmental Protection Agency [redacted]

**#13**

Presenter:
Russell
Erickson
U.S.Environmental Protection
Agency Mid-Continent Ecology
Division
[redacted]

Presenter Biography:
[redacted]

Title:
"Use of toxicity models in aquatic life criteria to address different levels of effect as a function of exposure concentration and time"

Abstract:
Aquatic risk assessments, including U.S.EPA aquatic life water quality criteria (ALWQC), are often limited regarding how well the likelihood and magnitude of toxic effects are quantified for a species. Generally, only a single level of effect is considered, rather than addressing how effects vary with
concentration. For example, U.S.EPA criteria maximum concentrations (CMCs) are based on acute toxicity tests characterized just by median lethal concentrations (LC50s). These CMCs do incorporate an extrapolation factor to represent a level of mortality lower than 50%, but the level of effect is then not only single but less clearly defined. Furthermore, these LC50s represent effects due to constant concentrations over fixed durations, with no assessment of how effects vary across different exposure time-series. This temporal variability issue is addressed by requiring CMCs to be satisfied over a specified averaging period, but this adds uncertainty regarding the actual level of effect represented by the criteria for the various exposure-time series to which they might be applied. Although this framework still provides a useful estimation of exposures of concern, better risk characterization should be possible based on considerable work over the 30 years since the issuance of the ALWQC guidelines to better understand and to model the magnitude of toxic effects as a function of exposure concentration and time. These various efforts will be reviewed and their applicability to ALWQC will be

discussed.

Russell Erickson
Mid-Continent Ecology Division
National Health and Environmental Effects Research Laboratory
U.S. Environmental Protection
Agency [redacted]

**#15**

Presenter:
Robert W. Gensemer,
Ph.D. GEI
Consultants, Inc.
[redacted]

Co-Authors:
John Gondek, Carrie A. Claytor, Shaun Roark, and Steven P. Canton, GEI Consultants, and Robert Dwyer, International Copper Association.

Presenter Biography:
[redacted]

Presentation Title:
Using the Biotic Ligand Model and Fixed Monitoring Benchmarks for resolving both spatial and temporal variability in setting aquatic life criteria for copper.

Abstract:
One of the central challenges for implementation of aquatic life criteria that are derived on the basis of water quality characteristics (e.g., water hardness, pH, or dissolved organic carbon (DOC) concentrations) is that concentrations of these characteristics vary over both time and space. Therefore, tools are needed to help guide the development of ambient water quality criteria that are sufficiently protective of aquatic life in this face of this variability.

The United States (US) Environmental Protection Agency's (EPA) Fixed Monitoring Benchmark (FMB) method is a probabilistic model that was developed to resolve the time-variability in Biotic Ligand Model (BLM)-derived instantaneous water quality criteria (IWQC), resulting in a single fixed regulatory value. However, in order to develop protective criteria for a regulated waterbody, data covering both multiple time periods as well as multiple sampling stations are often needed to fully characterize water quality and, therefore, derive protective criteria. Here we present the strengths and limitations of using the FMB to derive site-specific copper criteria with data from both multiple time periods as well as multiple sampling stations. These impressions are based on our recent experiences in Colorado, where we derived FMBs for multiple sites along Big Thompson River using 10 years of data. During this process, a number of issues were raised between stakeholders and regulatory agencies with respect to how one should compare BLM-based metrics between sampling stations and whether criteria generated using data aggregated across multiple sites using the FMB would be protective of aquatic uses at each individual site. We suggest that updates to EPA's guidelines for derivation of aquatic life criteria should incorporate the use of these probabilistic tools to better assist regulatory agencies and stakeholders select criteria concentrations that are protective of aquatic life under conditions of both temporal and spatial variability in water quality conditions.

**#16**

Presenter: Jeffrey M. Giddings

Organization: Compliance Services International

[redacted]

Biography: [redacted]

Presentation Title: Perspectives on the Derivation of Aquatic Life Criteria for Pesticides

Authors: Jeffrey Giddings, Compliance Services International; Dwayne Moore, Intrinsik Environmental Sciences

ABSTRACT
Pesticides registered for use in the US are required to undergo rigorous toxicity testing with a range of aquatic plant and animal species. The resulting data are used in a comprehensive ecological risk assessment. An important component of the risk assessment process is determination of the concentrations of the pesticide which, if not exceeded, will protect aquatic life – analogous to Aquatic Life Criteria (ALC). In 2011-12, CropLife America (a not-for-profit trade association representing the nation's developers, manufacturers, formulators and distributors of products for agriculture and pest management in the US) issued a series of white papers and participated in public meetings addressing harmonization of assessment methods used by the EPA Office of Water for ALC derivation and by the

EPA Office of Pesticide Programs for pesticide registration. Using this and other recently generated information, the current presentation will review several topics that are key to derivation of ALCs for pesticides:

(1) Protection goals that reflect the differences in ecological functions, sensitivity, and recovery potential of plant and animal populations and communities;
(2) Selection of toxicity data for ALC derivation, including consideration of data quality, species selection, measured endpoints, and test methodologies;
(3) Statistical methods for analysis of Species Sensitivity Distributions (SSDs), especially fitting distribution models to estimate concentrations likely to affect small fractions (e.g., $5^{th}$ or $20^{th}$ percentiles) of species;
(4) Derivation of Acute-to-Chronic Ratios (ACRs), including selection of toxicity endpoints, criteria for comparison of acute and chronic endpoints for single species, and development of default ACRs for use when data are limited;
(5) Concentration averaging periods and return frequencies that are consistent with protection goals for plants and animals;
(6) Use of mesocosm data to inform or adjust ALCs derived using laboratory toxicity data; and
(7) Triggers and criteria for adjusting ALCs if necessary to protect key species and ecological functions.

For each of the above topics, we will make recommendations on approaches and methods that may be used in derivation of ALCs for pesticides and possible other chemical stressors.

**#18**

Miranda Henning
Ramboll Environ
[redacted]

[redacted]

**Effective Weight-of-Evidence Approaches for Evaluating Ecotoxicity Studies**

The term “weight-of-evidence” is broadly used in diverse contexts to refer to methods for integrating the findings of multiple, sometimes conflicting studies in order to yield a single coherent conclusion that reflects the full body of relevant knowledge. Weight-of-evidence analysis offers important benefits and improvements to the development of aquatic life criteria, particularly with respect to identifying the ecotoxicity studies that form the basis for criteria. Professional judgment is an unavoidable element in the selection of one or more studies as the basis for a given aquatic life criterion, but a strong weight-of- evidence methodology can ensure that the basis for that professional judgment is transparent. An important benefit of weight-of-evidence evaluations is the clear documentation of the strengths, weaknesses, and uncertainties that underpin each applicable study. After exploring some of the history and benefits of weight-of-evidence approaches, this presentation examines both common and unique traits of several weight-of-evidence approaches in order to identify best practices within the context of setting aquatic life criteria. The approach advocated is modeled on that developed by the Massachusetts Weight-of-Evidence workgroup nearly 20 years ago, which employs Multicriteria Decision Analysis. The approach involves first assigning each available study a score that reflects its scientific strength, identifying the chronic value supported by each study, and examining the concordance among studies in scientific strength and chronic values. More specifically, scores are first assigned to each applicable study based on attributes of scientific strength such as study design, data quality, sensitivity/specificity, confounding factors, and statistical power. Second, each study is judged relative to each attribute and scored according to a rubric developed *a priori*. Because the attributes for judging study quality are not necessarily equally important, they may be weighted based on a survey of expert opinions. Weighted averages are calculated for each study in order to allow rank ordering of the studies. Each study’s chronic value is then plotted graphically, using symbols of varying size to reflect the studies’ scores for scientific strength. The plotted final chronic values are then examined for concordance or contrast in order to judge the range of final chronic values supported by the strongest studies. The calculated mean chronic value may be weighted by each study's scores, so that the mean is most strongly influenced by the strongest studies and least influenced by the weakest studies. The use of this quantifiable weight-of- evidence approach will be demonstrated using an example dataset.

**#20**

**Name of Presenter:** Carlie A. LaLone

**Organization:** U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Laboratory, Mid-continent Ecology Division

[redacted]

**Presenter Biography:** [redacted]

**Presentation Title:** Cross-species Evaluation of Molecular Target Sequence and Structural Conservation as a Line of Evidence for Identification of Susceptible Taxa to Inform Toxicity Testing

**Abstract:** The 1985 U.S. Environmental Protection Agency Guidelines for Deriving Aquatic Life Criteria require acute and chronic toxicity testing with a fixed list of taxa that cover a broad spectrum of aquatic organisms from vertebrate, invertebrate, and plant families. In considering revisions to the Guidelines, there is interest in employing new technologies that could inform chemical-specific selection of test species based on scientific evidence for susceptibility as a means to eliminate unnecessary animal testing, reduce cost, and gain the most applicable data for decision making. Further, when toxicity testing is completed on selected organisms, advances in methods for species extrapolation could be implemented to gain a greater understanding of how representative the data are of other species from the same taxonomic group. A currently available computational method for informing such considerations is the U.S Environmental Protection Agency's Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool. Developed as a product under the Chemical Safety for Sustainability Research Program, SeqAPASS facilitates cross species comparisons of chemical molecular targets based on evaluation of protein sequence and structure. The assumption underlying this approach is that if a chemical is known to act on a specific protein to initiate toxicity in one species, conservation of that protein in another species provides a line of evidence that the second species could also be susceptible to a given chemical. Depending on the degree of characterization of the selected protein and the existing knowledge about the chemical-protein interaction, the SeqAPASS tool derives susceptibility predictions by calculating percent similarities from alignments of primary amino acid sequences, functional domains (e.g., ligand-binding domain), and individual amino acid residue(s) important for maintaining protein structure or for direct interaction with the chemical. Such data can be used to identify both species that are likely susceptible and those

unlikely susceptible to a given chemical, therefore aiding in selection of appropriate test species. Another application for the data would be in understanding how broadly toxicity test data derived from one species may be extrapolated to others. In short, evaluation of cross species protein and structural conservation using SeqAPASS provides a rapid and cost effective method for deriving a line of evidence for chemical-specific testing strategies that may be of utility in the revised Aquatic Life Criteria Guidelines. *The contents of this abstract neither constitute nor reflect official US EPA policy.*

**#22**

**Presenter:**
**Samuel N**
**Luoma**

**Affiliation:**
**John Muir Institute of the Environment**
**University of California, Davis**
[redacted]

**Mailing address:**
[redacted]

**Telephone:**
[redacted]

**Email:**
[redacted]

**Biographical Statement: Samuel N Luoma**

[redacted]

**Abstract:**

**Criteria for Determining the Value of Bringing More "Eco" into Ecotoxicological Guidance for the Development of Water Quality Guidelines.**

Existing USEPA criteria for most metals and metalloids in freshwater and salt water are built from a scientific base of robust toxicological approaches that were initially developed before 1985 and refined thereafter. It is often stated that the guidelines as constituted are over- protective and that greater flexibility is needed to deal with site-specific conditions. As a result new knowledge of aquatic geochemistry as it affects metal toxicity has been a major focus of research and increasingly incorporated into at least some of the criteria. A robust "parallel literature" also developed over the last twenty years that illustrates biological and ecological shortcomings in the existing criteria. Unlike geochemistry, nearly all of this literature is formally excluded from consideration by the minimum data requirements (Luoma and Rainbow 2015, Chapter 15). Here we propose that a systematic assessment is needed to determine whether a more comprehensive consideration of biological and ecological understanding could help better achieve the flexibility necessary to deal the site-specific aspects of metal risks as well as more efficiently and effectively protect the environment. Freshwater and saltwater criteria for cadmium and copper will be evaluated as an example to demonstrate some guidelines useful in evaluating the compatibility of the existing criteria with the more comprehensive literature. Questions that will be addressed in the evaluation of the criteria include:

- Was chronic toxicity, relevant to population processes, adequately addressed? If not what are the implications?
- Is the justification for species sensitivity robust and is it consistent with field knowledge of such sensitivities?
- Are the criteria credible when compared with dissolved concentrations known from modern geochemical analyses of contaminated and uncontaminated waters?
- How do the criteria compare to thresholds of effects seen in complex toxicological studies of effects, like mesocosm or life cycle tests?
- How do the criteria compare to studies that include dietary exposure?
- Are the guidelines justifiable based upon field observations that meet criteria for appropriate effects analysis? How do they compare to robust field or integrated laboratory-field observations?

A comprehensive evaluation following modified guidelines of this sort could identify criteria least compatible with the full literature and thus most in need of reform. Methodologies and models will be discussed that could be used in a new generation of guidelines that better links biology, ecology and ecotoxicology as well as field observations and laboratory assays.

References:

Luoma SN and Rainbow PS. 2008. Metal Contamination in Aquatic Environments: Science and Lateral Management. Cambridge University Press. Cambridge UK. 573pp

# 23

Abstract for Invited Expert Meeting on Revising USEPA's Guidelines for Deriving Aquatic Life Criteria, September 14-16, 2015

Improving Uncertainty Characterization in USEPA's Guidelines for Deriving Aquatic Life Criteria Using Decision Contexts

Douglas B. McLaughlin

The development of numeric criteria for the protection of aquatic life plays an important role in the management of risks to aquatic ecosystems. Numeric criteria can effectively summarize considerable scientific complexities, providing a management tool upon which useful regulatory decisions may be based. The 1985 Guidelines typically yield single criterion values (e.g., Final Acute Value, Final Chronic Value, Criterion Maximum Concentration, Criterion Continuous Concentration), but provide little guidance on how the uncertainty in these values might be evaluated or expressed. Characterizing uncertainty is a critical scientific endeavor, and has previously been identified as a priority for revisions to the 1985 Guidelines. A large amount of published information in water resources and other fields is available to support guidelines revisions that address these limitations. For example, (McLaughlin and Jain 2011) use Monte Carlo methods applied to the acute criterion derivation equations in the 1985 Guidelines to characterize uncertainties in acute criteria derived from variation observed in laboratory results and limitations in the amount and type of available toxicity data.

However, revised guidelines might also include specific guidance on ways to characterize uncertainty that can be more easily understood by stakeholders having a range of technical expertise. As described by Preston et al. (2004), successfully communicating or "translating" science, including uncertainty, to many stakeholder groups is critical to the effective management of ecosystem risks. Numeric criteria are ultimately used as a basis for making decisions about the attainment of water quality standards, the need for total maximum daily load studies, allocations, and load restrictions, NPDES permit limitations, and others. Each type of decision provides a context for considering the importance of scientific uncertainty when managing aquatic systems with the aid of aquatic life criteria. For example, McLaughlin (2012a, 2012b, 2014) describes several approaches for characterizing uncertainties in numeric water quality criteria as probabilities of correct and incorrect water quality attainment decisions. McLaughlin (2015) illustrates how the agreement between observed and expected effect concentrations in Biotic Ligand Model validation data, typically characterized as "within a factor of 2", also can be presented as the probability of correct and incorrect "excess risk" decisions using risk ratios such as toxic units.

This presentation will review these and other related studies, providing examples of how scientific uncertainties in numeric criteria can be characterized with available statistical and probabilistic tools, and how the level of risk associated with criteria can be estimated and communicated to a broad audience using an appropriate decision context. Considering these concepts when revising the aquatic life derivation guidelines will improve the transparency of aquatic life criteria across many stakeholder groups, provide a more complete assessment of the state of the science supporting criteria derivation,

more clearly establish research priorities for improving the predictive value of aquatic life criteria, and provide a better basis for managing risks using criteria science.

REFERENCES

Preston, HR, J Culp, T DeMoss, R Freydenborg, EE Kenaga, CA Pittenger, N Schofield. 2004. Translating ecological science. in Barbour MT, SB Norton, HR Preston, K Thornton, editors. 2004. Ecological assessment of aquatic resources: Linking science to decision-making. Pensacola (FL): SETAC Press. 272 p.

McLaughlin, DB 2015. Assessing the fit of biotic ligand model validation data in a risk management decision context. Integrated Environmental Assessment and Management, DOI: 10.1002/ieam.1634

McLaughlin, DB 2014. Maximizing the accuracy of field - derived numeric nutrient criteria in water quality regulations. Integrated Environmental Assessment and Management 10 (1): 133-137.

McLaughlin, DB 2012a. Assessing the predictive performance of risk-based water quality criteria using decision error estimates from ROC analysis. Integrated Environmental Assessment and Management. 8(4): 674-684.

McLaughlin, DB 2012b. Estimating the designated use attainment decision error rates of USEPA's proposed numeric total phosphorus criteria for Florida colored lakes. Integrated Environmental Assessment and Management 8(1):167-174.

McLaughlin, DB and V Jain. 2011. Using Monte Carlo Analysis to characterize the uncertainty in final acute values derived from aquatic toxicity data. Integrated Environmental Assessment and Management 7(2): 269-279.

Biography of Douglas B. McLaughlin, Ph.D

[redacted]

**#25**

**Invited expert meeting on revising U.S. EPA's guidelines for deriving aquatic life criteria, September 14-16, 2015, Arlington, VA**

**Name of presenter, organization, address, phone, email:**
Kathleen Patnode, U.S. Fish and Wildlife Service, [redacted]

**Presenter biography:**
[redacted]

**Presentation title:**
Incorporation of field or meso/microcosm data to validate criteria in watersheds supporting federally listed species

Kathleen Patnode, Lora Zimmerman, and Robert Anderson (USFWS, [redacted] and Amy Bergdale (USEPA, [redacted]

**Abstract:**
Federally listed aquatic species are rarely used in laboratory testing due to difficulty in obtaining and culturing these species, permitting requirements, and prohibitions on sacrificing organisms that are exceptionally rare. As a result, water quality criteria (WQC) derivations do not include data for many of these species, and the resulting discharge criteria in NPDES permits may not be protective of these animals.

In 2012, U.S. Fish and Wildlife Service (The Service) and U.S. Geological Survey conducted an *in-situ* study using northern riffleshell (*Epioblasma torulosa rangiana*), a federally listed freshwater mussel, to evaluate the protectiveness of a NPDES-permit for an industrial discharge in the Allegheny River. The results of this study demonstrate that federally listed freshwater mussels may experience lethal and sublethal effects at permitted discharges which constitutes unauthorized "take" under the Endangered Species Act. Several factors contributed to the observed toxic effect: failure of the state agency to acknowledge the presence of a federally listed species, high sensitivity of freshwater mussels to ionic contaminants, discharge of chloride in exceedance of the no adverse effect concentration for this species, characterization of the effluent insufficient to derive necessary permit limits, additivity or synergism of multiple ionic contaminants, allowance for a mixing zone, and reliance on limited duration toxicity testing which is not representative of river exposures.

While some of these factors are specific to this site, several have ramifications for WQC derivation and implementation. The high sensitivity of freshwater mussels to ions such as ammonia and chloride has been demonstrated in laboratory testing. We support the USGS recommendations that toxicity testing in at least one species of Unionidae be included in WQC development. However, laboratory testing fails to replicate the conditions experienced by these species in the field – exposure to mixtures of these and other contaminants for months, rather than days. Thus, the industrial discharge that we studied could have resulted in take of northern riffleshell without exceedances of the current chloride WQC or the revised ammonia WQC. To perform tests of mixtures for extended duration in the laboratory is logistically and economically impractical. We contend that our study and other site-specific caged mussel investigations demonstrate the value of criteria validation via *in-situ* testing, particularly in surface waters supporting or capable of supporting federally listed mussel species. As surface water chemistry varies regionally and substantially influences toxicity of some contaminants, *in-situ* validation is applicable at the watershed level where federally listed species occur. Such validations could be incorporated into the triennial review process of each state. Regional EPA offices could oversee validations to test extended single exposures, as well as common mixtures using existing NPDES permits and DMR data. The Service could provide technical expertise on mussel ecology and handling. These cooperative validations would ensure that state-promulgated water quality criteria are protective of federally listed species and are in compliance with all state and federal regulations.

**#31**

**Name of Presenter:** William Stubblefield
**Organization:** Oregon State University
**Address:** [redacted]
**Telephone:** [redacted]
**Email Address:** [redacted]
**Presenter Biography:**
[redacted]

**Presentation Title**: Water Quality Criteria/Standards: International Approaches for Derivation
**Abstract:**
Central to the US Clean Water Act is the development of numeric Ambient Water Quality Criteria [AWQC] for individual compounds or elements that, when not exceeded, are believed to be protective of aquatic organisms and aquatic ecosystems from adverse effects. These values are intended to be conservative in nature and applicable to most, if not all, waters of the US. The criteria derivation procedure was first published in 1985 (Stephan et al., 1985) and has been in place unchanged since that time. It was recognized that in some instances "National" criteria may be either over-protective or, in some cases, under-protective of certain aquatic systems and in those cases guidelines for the site-specific modification of national criteria were provided (Carlson et al., 1984). Since the development of the USEPA criteria derivation approach, there have been many advances in the science of environmental toxicology and a number of lessons learned regarding the implementation of criteria/standards. In addition,

several other jurisdictions (e.g., European Union, Canada, New Zealand/Australia) have developed their own approaches for deriving chemical-specific numerical standards. While independently developed, all of the approaches share commonalities but have important differences. In general the approaches tend to be based on available laboratory-generated toxicity data for a variety of aquatic species, and these data are extrapolated or interpolated in different ways to derive a criterion/standard. Requirements for the species to be represented in the database, acceptable test methods, and data quality requirements differ among regulatory jurisdictions. Nonetheless, all of these approaches assume that criteria derived from a subset of species will be acceptably protective of the majority of species, both tested and untested, and that species protection confers ecosystem protection. For example, in the U.S., aquatic life criteria are typically derived from a minimum database that includes acute toxicity tests for at least 8 representative genera and chronic toxicity tests for at least 3 genera. Europe, on the other hand, requires data for 8 representative groups (although the groups differ from the USEPA) and they are primarily interested in chronic toxicity data. Similarities and differences in regulatory approaches will be compared and the pros and cons of the approaches discussed.

#32

**Invited Expert Meeting on Revising U.S. EPA's Guidelines for Deriving Aquatic Life Criteria**

**14-16 September 2015, Arlington, VA**

**Name of Presenter**: G. W. Suter

**Organization:** NCEA[redacted], ORD, EPA

**Address, City, State, Zip:** [redacted]

**Telephone:** [redacted]

**Email Address**: [redacted]

**Presenter Biography** (300 words max)

[redacted]

**Presentation Title:** Application of Weight of Evidence Methods to Water Quality Criteria

**Abstract** (500 words max)

When more than one piece of evidence is available to address a problem, it is appropriate to weigh the evidence. Weight of evidence (WoE) may be used to address multiple aspects of criteria development.
(1) Conventionally, the geometric means of multiple test results for a species are used to derive water quality criteria. However, one might quantitatively weight those values before averaging, as in meta analyses (e.g., inverse variance weighting). (2) One might also weight values in species sensitivity distributions (SSD) based on qualities such as relevance, representativeness, or reliability (e.g., use of good laboratory practices). (3) Like human health benchmarks, WoE may be used to determine what mode of action or route of exposure is applicable (e.g., is it an endocrine disrupter or is dietary exposure significant?) to an aquatic life criterion. (4) When deriving criteria using field data, weight of evidence is used to determine whether the exposure-

response relationship is causal and whether the models are significantly confounded. (5) WoE may also be used to determine whether a field-derived value for one region is applicable to another. (6) The ultimate application of weight of evidence would be to choose among or combine potential criteria derived by different methods (e.g., conventional toxicity tests, microcosms, mesocosms, field surveys, or models). (7) Finally, weight is the opposite of uncertainty. Therefore, qualitative weighting of a body of evidence, based on consideration of relevance and reliability, can be used as an expression of the confidence in a value or conclusion that goes beyond the scatter that is quantified by statistical uncertainty. Uses of WoE beyond geometric means will require research and demonstration projects. The views expressed in this abstract are those of the author and do not necessarily represent the views or policies of the U.S. EPA.

**Key Words:** weight of evidence, alternative approaches, statistics, uncertainty

**#33**

**Invited expert meeting on revising U.S. EPA's guidelines for deriving aquatic life criteria, September 14-16, 2015, Arlington, VA 22202**

**Name of presenter, organization, address, phone, email:**

Ning Wang, U.S. Geological Survey, [redacted]

**Presenter biography (300 word max):**

[redacted]

**Presentation title:**

Minimum data requirement for developing water quality criteria: Use of toxicity data from under-represented organisms

Ning Wang, John M Besser, Chris G Ingersoll (USGS, Columbia, MO), Tom Augspurger (USFWS, [redacted]), Sandy Raimondo (USEPA, [redacted]), Ed J Hammer, Candice R Bauer (USEPA, [redacted] and Chris A Mebane (USGS, [redacted]

**Abstract (500 words max):**

USEPA ambient water quality criteria (WQC) are developed using a genus level sensitivity distribution approach, which is typically based on laboratory toxicity data from a suite of aquatic organisms that are assumed to represent the sensitivity of untested species (Minimum Data Requirement of 8 different families). Over the past 20 years, we developed, validated, and applied methods to estimate acute and chronic toxicity of contaminants with different toxic modes of action to more than 30 federally threatened or endangered (listed) species and more than 20 species of freshwater mussels, sculpins, and fairy shrimp typically under-represented in toxicity databases used to derive WQC. We compared the acute or chronic sensitivity of these

species to commonly tested fish (e.g., fathead minnow *Pimephales promelas* and rainbow trout *Oncorhynchus mykiss*) and aquatic invertebrates (e.g., cladoceran *Ceriodaphia dubia* and amphipod *Hyalella azteca*) that are often included in toxicity databases used for deriving WQC, and evaluated the protectiveness of WQC for listed and under-represented species. We also conducted mussel tests to evaluate influences of life stages (e.g., larvae vs juveniles) and chronic toxicity test duration (sodium chloride exposures of 4 vs 12 weeks) on mussel sensitivity, and determined the interspecies variability. In general, the listed or under-represented fish were equally or less sensitive compared to rainbow trout in acute exposures and were adequately protected by acute WQC when trout data are retained. However, chronic WQC were not always adequately protective, as demonstrated by chronic copper exposures with fountain darter *Etheostoma fonticola*, mottled sculpin *Cottus bairdii*, and white sturgeon *Acipenser transmontanus*. Freshwater mussels were generally more sensitive than commonly tested invertebrates or fish in acute and chronic exposures to metals (e.g., copper, nickel, and zinc), major ions (e.g., chloride, potassium, and sulfate), and ammonia, and WQC for most of these chemicals are not adequately protective of mussels. Based on a compiled acute toxicity database with 479 records for 49 freshwater species and 40 chemicals, 3 commonly tested mussel species represented other mussel species (acute values within a factor of 2) in most instances (69-79%) whereas cladocerans represented mussels for 35-36% of instances and fish represented mussels for only 10-25% of records. Results of mussel studies also indicate that various life stages of mussels had different sensitivity to some chemicals, and chronic effects of sodium chloride on mussel growth increased with extended exposure periods from 4 to 12 weeks. Improved WQC for freshwater organisms would be obtained if the Minimum Data Requirement for deriving WQC were updated to: (1) include native mussels as a required family; (2) use mussel growth data from 4- to 12-week toxicity tests to derive chronic WQC; (3) include sensitive species, even though some of them (e.g., rainbow trout) are not prevalent in some waters, to better represent the sensitivity of under-represented species, (4) include nontraditional test species and associated nonstandard test methods to improve the protectiveness of WQC; and (5) develop water quality guidelines for chemicals of concern when toxicity data for taxa representing less than 8 families are available for WQC derivation.

**#38**

Frank S Dillon
Principal
Technologist

CH2M HILL
[redacted]

[redacted]

[redacted]

[redacted]

- [redacted]

[redacted]

**Statistical and Biological Interpretation of Test Endpoints Relative to Unacceptable Risk Levels**

Current guidance specifies that chronic criteria can be derived through a series of chronic tests with a range of species or through the development of acute-to-chronic ratios. There is the potential for significant variation in ACRs depending on a number of test factors and statistical analysis approaches. This will be illustrated via a series of tests that served as the basis for developing recommended site- specific criteria for metals. In order to estimate ACRs relevant to a coldwater stream community, we exposed rainbow trout (Oncorhynchus mykiss) to cadmium (Cd), lead (Pb), and zinc (Zn) in 96-h acute and 60+ day early-life stage (ELS) exposures. We also tested the acute and sublethal responses of a mayfly (Baetis tricaudatus) and a midge (Chironomus dilutus, formerly C. tentans) with Pb. We examined the statistical interpretation of test endpoints and the acute-to-chronic ratio concept. In the tests we reviewed for comparison with our results, hypothesis testing was by far the most commonly used statistical analysis approach to establish chronic values and a few studies used no statistical analyses whatsoever. However, regression techniques have a number of advantages over hypothesis testing for making point estimates of low toxic effects, and so where possible we calculated comparable ECp values from the previous reports. In our tests and in most of the tests we reviewed and calculated ECp values from, NOECs were about as high or higher than EC10 values. In other words, NOECs corresponded with low-effect concentrations, not no-effect concentrations. Increasing the number of control replicates by 2 to 3× decreased the minimum detectable differences by almost half. Pb ACR estimates mostly increased with increasing acute resistance of the organisms (rainbow trout ACRs $<\approx$ mayfly $<$ Chironomus). The choice of test endpoint and statistical analysis influenced ACR estimates by up to a factor of four.

**Declined**

#06

**Name:** [redacted]

**Organization:** [redacted]

**Address:** [redacted]

**Telephone:** [redacted]

**Email Address**: [redacted]

**Presenter Biography:**

[redacted]

**Presentation Title:**

Aquatic Life Criteria should consider biotic interactions with sediments, hyporheic zones and coexisting stressors in a weight-of-evidence based approach

**Abstract:**

Aquatic ecosystems are integrated functioning systems and should not be artificially separated into surface water, sediment, groundwater and terrestrial (riparian) compartments as in regulatory programs. Chemical inputs often travel through all ecosystem compartments which alter their fate, bioavailability and exposures to organisms dominating those compartments. For example, some pelagic organisms (e.g., *Daphnia magna* and *Pimephales promelas*) feed on sediment surfaces and are also exposed to sediment-bound chemicals via resuspension, bioturbation, upwelling from groundwaters/hyporeheic zones or emerging from sediments via other advective/diffusion processes. Nevertheless - aquatic life criteria are based on dissolved water concentrations only. In addition, all human-dominated waterways have multiple stressors to which biota are exposed, interacting directly and indirectly with each other and resident biota. Decades of studies have shown how to improve the accuracy of ecological risk assessment methods by encompassing spatial and temporal variability, *in-situ* exposures and "smartly" designed weight-of-evidence based assessments using tried-and-true methods. More recently this approach has been used to better understand the ecological significance of pharmaceuticals and personal care product releases from WWTPs and farming operations. Findings tend to agree with EPA 303(d) list

rankings showing habitat components, nutrients and flow dominate as stressors and must be dealt with before the goals of the Clean Water Act (i.e., restoring the physical, biological and chemical integrity of the Nation's waters) can be attained.

[redacted]

**#07**

Presenter:

[redacted]

Co-Authors:

[redacted]

Presenter Biography:

[redacted]

Presentation Title:

Use of the recalculation procedure to develop site-specific standards for arid West effluent-dependent waters.

Abstract:

Although AWQC developed using the 1985 Guidelines are intended to protect all but a small percentage of the most sensitive aquatic species nation-wide, they may not always represent the contaminant sensitivity of species resident to a particular site. We present a revised method for developing ambient water quality criteria specific to effluent-dependent, non-perennial streams in the arid West, based on EPA's recalculation procedure for derivation of site-specific criteria. We also recommend several modifications to the recalculation procedure that could also be broadly applicable outside the arid West. First, we examine "resident" vs. "transient" species when evaluating "expected condition," accompanied by a refined step-wise deletion process with the goal of generating a site-specific toxicity database more representative of the species that occur or could occur. Second, we recommend redefining the minimum data requirements (MDRs) for the toxicity database, with a revised eight-family rule intended for protection of

warm water aquatic communities residing in arid West effluent-dependent streams. Third, we also recommend calculation of criteria using the geometric mean of species mean acute and chronic values (SMAVs and SMCVs) rather than genus mean acute and chronic values (GMAVs and GMCVs) to help increase the database size and resolve potential sample size effects without impacting the protectiveness of the resulting criteria. This change also resolves issues that can arise when species within a genus have widely different sensitivities to a toxicant, which can get lost when combining into a GMAV. Using this process case studies were conducted with aluminum, ammonia, copper, diazinon, and zinc, with resulting site-and/or-region-specific criteria often significantly different for some, but not all of these parameters. Based on our analyses, the modified recalculation procedure can be a useful tool when applied to arid West streams and the modifications could be considered more broadly to perhaps develop national use-based (i.e., “cold”, “warm”, etc.) criteria.

[redacted]

**#14**

Abstract Submission: Invited Expert Meeting on Revising US EPA's Guidelines for Deriving Aquatic Life Criteria

Topic: Approaches for adjusting criteria based on physicochemical factors

**Title: Derivation and implementation of a bioavailability based Environmental Quality Standard for nickel under the European Union's Water Framework Directive**

Authors: [redacted]

[redacted]

[redacted]

The toxicity of nickel in freshwaters depends upon its bioavailability under the local water chemistry conditions. In Europe, nickel is classified as a Priority Substance, meaning that all 27 European Union (EU) Member States need to comply with a single Environmental Quality Standard (EQS) under the Water Framework Directive (2000/60/EC). Guidance for the determination of EQS values allows for the use of advances in risk assessment tools like Species Sensitivity Distributions and bioavailability normalization. The guidance also addresses the use of bioavailability normalization of ecotoxicity data, and specifically calls for the development of chronic models for three trophic levels, i.e., algae/plants, invertebrates, and fish. To this end, chronic Ni bioavailability models have been developed and validated for four different aquatic species (an alga, two invertebrates, and a fish). Briefly, these models demonstrated that Ni toxicity is affected by hardness, pH, and dissolved organic carbon (DOC) for all three trophic levels. The models have been shown to be applicable to other species for which no species-specific models exist, demonstrating the universality of the bioavailability concept. This is important considering that the chronic Ni database is comprised of 31 species of algae, vascular plants, invertebrates, fish, and amphibians. The bioavailability models can therefore be used to predict the effect of water chemistry on nickel toxicity to each individual species for which toxicity information is

available. To derive a nickel EQS, nickel bioavailability was assessed for EU surface waters with varying water chemistries (varying pH, DOC, hardness) for which sufficient water chemistry data were available. The results of this evaluation were used to identify the most sensitive EU surface waters for nickel toxicity to derive a reference EQS to ensure protection of 95% of the waters within the most sensitive European region from nickel toxicity. The reference EQS was determined by normalizing all chronic ecotoxicity data (i.e., EC10s) to the conditions of

[redacted]

the identified sensitive waters (DOC = 2 mg/L, pH = 8.2, Ca = 40 mg/L). The normalized ecotoxicity data were used to populate a Species Sensitivity Distribution to calculate an HC5, which for the most sensitive waters was 4 µg Ni/L. The bioavailable nickel EQS is implemented within a tiered compliance assessment approach in which the first tier is a direct comparison of the local dissolved nickel concentration with the reference EQS of 4 µg Ni/L. Under situations where dissolved Ni is > 4 µg Ni/L, subsequent tiers allow for correction of the nickel exposure to account for bioavailability based on local water chemistry conditions. A user-friendly tool (bio-met.net) that is based on full BLM normalization has been developed to facilitate the implementation of the tiered approach.
Bio-met uses a reduced set of input parameters (pH, DOC, and Ca), which increases its utility among EU Member States that have limited monitoring capabilities.

Presenting Author:

[redacted]

*Biography attached*

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

3

**#17**

ABSTRACT

## Modifying AWQC Based on Site Specific Conditions – What's Going to Guide the Process?

[redacted]

Section 304(a) of the Clean Water Act authorizes EPA to develop and publish national ambient water quality criteria (AWQC) based on scientific information. These national criteria are issued periodically to the states as guidance for use in developing their own criteria. EPA has long recognized in its regulations, however, that the national recommended water quality criteria published as guidance may be modified "to reflect site specific conditions." 40 CFR 131.11(b)(1). As EPA looks forward toward revising the AWQC derivation process, this process must continue to allow for the ability for AWQC to be modified to reflect site specific conditions. Several methods are currently available for modifying AWQC based on site specific conditions. However, the ability to effectively apply these methods can be significantly affected by the guidance that accompanies the method. This presentation will contrast the recent application of two existing methods, the water-effects ratio and the copper biotic ligand model, to illustrate the importance of having adequate guidance that clearly provides the information needed by both the regulated community and regulators to perform and analyze site- specific studies for permitting purposes.

[redacted]

[redacted] [redacted] [redacted] [redacted]

**#19**

Presenter: [redacted]

Organisation: [redacted]

Zealand

Address: [redacted]

Telephone: [redacted]

Email: [redacted]

Topic: Selecting and prioritizing pollutants of concern

Title: Nitrate toxicity to freshwater aquatic species: updating the Australian and New Zealand guidelines and standards

Authors: [redacted]

**Abstract**

Nitrate occurs naturally in the environment and is produced and consumed through the processes of the nitrogen cycle, and anthropogenically produced for agricultural use as a fertilizer. The major anthropogenic sources of nitrate to surface waters are agricultural runoff, municipal and industrial wastewaters, urban runoff and groundwater inputs. Nitrate concentrations are an important indicator of agricultural enrichment and ecological health – affecting both surface water and groundwater environments. As such, they form an important component of the management of freshwaters, requiring robust guideline values to support environmental planning and management.

The current revision of the Australian and New Zealand (ANZECC) water quality guidelines derives nitrate guideline values (GVs) for a range of protection levels. The New Zealand National Policy Statement for Freshwater Management (NPS-FM) has recently legislated nitrate standards for all river environments – incorporating a national "bottom line" which all waters are expected to achieve. These standards use a two-number guideline and management framework based on the ANZECC risk-based methodology to provide various levels of ecosystem protection. The legislation requires the compulsory application of this framework for water classification.

Various studies have demonstrated a reduced sensitivity to nitrate for both fish and invertebrates as water hardness increases. Increased chloride concentrations have

been demonstrated to reduce nitrite toxicity and a similar mechanism may occur for nitrate. The revised ANZECC guidelines and New Zealand's NPS-FM standards would be considered conservative as the most sensitive species in the database used to derive the nitrate guidelines is for long-term tests undertaken in very low hardness water – which would be expected to result in the highest nitrate toxicity. Conversely, application of these guidelines to differing water quality conditions may result in excessive protection. The application of a hardness-based correction algorithm based on existing studies indicates that toxicity mediation may be substantial and that adjusted criteria should be developed.

Nitrate toxicity has the potential for adverse effects on sensitive and on rare and endangered species, such as freshwater mussels. Nitrate should be considered a high priority for the development of regulatory criteria because of its widespread occurrence and often increasing concentration trends in freshwaters.

**Presenter Biography**

[redacted]

[redacted]

[redacted]

[redacted]

**#21**

[redacted]

Bio:
[redacted]

Abstract: Considerations for Validating Toxicity Test Data for Freshwater Mussels as used in Development of Aquatic Life Criteria

Aquatic life criteria (ALC) are developed to protect sensitive aquatic taxa. In the last decade, freshwater mussels (Uninoida, primarily the families Unionidae and Margaritiferidae) have been identified as an important faunal class that can demonstrate high sensitivity to commonly encountered toxicants in the aquatic environment. Ex-situ propagation programs of endangered and key mussel community species have been developed for various river recovery plans. In 2005, ASTM Guidance E2455-05 *Standard Guide for Conducting Laboratory Tests with Freshwater Mussels* was approved as an acceptable method for conducting toxicity tests with freshwater mussels. The methods established therein primarily responded to research that reported high sensitivity of mussels to ammonia, and the need to update the ammonia ALC for protection of important faunal group. Since 2005, many researchers have conducted toxicity tests with mussels for purposes other than ALC development, and have demonstrated test performance criteria that place the results of mussel tests in parity with the more commonly tested USEPA and OECD test species. However, the 1985 USEPA ALC Guidelines do not reflect the rigor of data quality parameters of the current USEPA and OECD methodologies. Specific to mussel testing, the following areas of consideration will be discussed for improving the quality of mussel toxicity tests for use in ALC derivation:

- Test endpoints such as EC50, LC50, and NOEC values as related to age groups

(glochidia, drop- offs, two-month old test organisms)

- Test endpoints as related to observational limitations (foot versus cilia motion, heartbeat, and magnification limitations), and recovery responses
- Application of test dose-response and precision guidelines
- Species selection and culture history
- Water-only versus substrate additions (primarily chronic tests)
- Dilution water quality and feeding (primarily chronic tests).

**#24**

Presenter: [redacted]

[redacted]

[redacted]

Phone: [redacted]

Email: [redacted]

Bio: [redacted]

Title: Incorporating Weight-of-Evidence and Biological Endpoints in a Water Quality Standard for Nutrients

**Abstract**

A collective body of work exists that effectively describes the relationships between nutrients, enrichment, and biological condition in running waters such that numeric endpoints can be identified for use in management. However, unlike a criterion derived from toxicity endpoints wherein the risk to aquatic life from exceeding the criterion value is fairly certain, the risk to aquatic life from exceeding a numeric endpoint for nutrients is far less certain. The uncertainty surrounding nutrient endpoints holds two main consequences for application in a water quality standard. First, additional information from a waterbody must be gathered and applied as weight-of-evidence to assess risk and inform management. And secondly, once a numeric target is chosen for a given waterbody, assessing compliance with the target value must be informed by local conditions. Here, we cover aspects of the first consequence; specifically, interpreting response indicators in a weight-of-evidence approach, and choosing appropriate target values based on the weight-of-evidence.

---

[1] Contact information and bio [redacted]

[redacted]

[redacted]

**#26**

Presenter: [redacted]

[redacted] Organisation [redacted]

Address: [redacted]

UK Telephone: [redacted]

Email: [redacted]

Topic: Approaches for adjusting criteria based on physicochemical factors

Title: Validation of the nickel biotic ligand model for locally relevant species in Australian freshwaters

Authors: [redacted]

[redacted]

**Abstract**

Biotic ligand models (BLM) for nickel have been developed based predominantly on tests performed on European and North American species and according to the water chemistry conditions which are prevalent in those regions. To apply the nickel BLM in Australia some key uncertainties need to be addressed, including the applicability of the models to Australian surface waters and a lack of chronic toxicity data from tropical species. The geochemistry of Australian freshwaters was assessed showing that many waters have relatively low water hardness (25% of the dataset below 6 mg $CaCO_3$ $l^{-1}$) and a different Ca to Mg ratio compared with many European and North American waters. The hardness of a large proportion of Australian freshwaters fall outside of the application boundaries of the existing NiBLM.

Therefore, ecotoxicity testing was performed to validate the performance of the models on five species, a fish, a cnidarian, a crustacean, an alga and an aquatic plant, in five Ni spiked field collected natural waters that varied with respect to water hardness, pH, and dissolved organic

carbon. At a broad level, Ni ecotoxicity varied considerably among the different test waters, which is consistent with previous studies on effects of Ni performed in natural waters. Overall, no single water chemistry parameter was able to indicate the trends in toxicity to all of the test species. Predictions of the toxicity using the unmodified EU NiBLM were relatively poor for the soft waters. An additional programme of testing was undertaken using the tropical green Hydra in order to assess the predictability of the existing NiBLMs for this species in extremely soft to moderately soft waters. Specifically, tests were performed in waters ranging in hardness from 1.1 to 50 mg CaCO3 $l^{-1}$, which covered equally broad ranges of Ca (0.05 to 10.0 mg Ca $l^{-1}$) and Mg (0.2 to 7.0 mg Mg $l^{-1}$). The testing programme revealed that the effect of Ca and Mg in competing with nickel for binding to the biotic ligand was substantially higher in very soft water (hardness < 10 mg CaCO3 $l^{-1}$) than the effect under higher water hardness conditions.

Modifications were made to the NiBLM to account for the softer waters encountered in Australia and the more important competitive effect of Ca and Mg on nickel toxicity in extremely soft waters. Nickel binding to DOC in the test waters was taken into account in the same manner as for the European model, using WHAM6. The modified model was better able to predict the toxicity of nickel to the test species in the validation studies in field collected waters than the existing model which is currently applied in Europe. This experience indicates the need to validate models within the ranges of geochemical parameters that will be expected in the geographical region of choice. It also confirms the importance of bioavailability normalization which takes account of all of the important factors when setting environmental quality standards for metals like Ni.

**Presenter Biography**

[redacted]

[redacted]

[redacted]

#27

Abstract for 2015 Invited Expert Meeting on Guidelines Revisions

July 2, 2015

## 1 Presenter Information

[redacted]

## 2 Presenter biography

[redacted]

[redacted]

[redacted], [redacted]

ers a wide range of environmental and ecological topics, including modeling phosphorus retention in the Everglades wetlands, detecting and quantifying ecological thresholds, water quality standard compliance study, eff of urbanization on stream ecosystem, and various ecological topics. His research emphasizes the development and adaptation of statistical methods suitable for environmental problems. He developed the Bayesian hierarchical model for EPA's 6-year review of drinking water standard compliance status, introduced the hockey stick modell for developing numerical phosphorus criterion, applied the seasonal trend analysis using loess for assessing long term trends in nutrient concentrations in the Neuse River basin, developed the Bayesian SPARROW model, introduced the multilevel models to study the eff of urbanization on stream ecosystem, introduced the use of the change point model for nutrient criterion development, and the use of advanced statistical methods for analyzing species compositional data. In 2015, he published three articles discussing statistical issues in environmental standard compliance assessment.

## 3 Presentation Title

Setting and Assessing the Compliance of Numerical Water Quality Criteria – Statistical Considerations

## 4 Abstract

Statistical issues frequently receive less attention in setting and assessing the compliance of a numerical criterion. In setting a numeric criterion, we either develop a dose-response model for a toxic material or estimate the reference probability distribution (also a model) of a water quality constituent (e.g., nutrient) concentration. The resulting models are the basis for setting the criterion, a process of decision-making under uncertainty. In assessing the compliance of a criterion, we compare the distribution of a pollutant's concentration to the criterion. But we must estimate the distribution using monitoring data with a small sample size. It is also a decision-making under uncertainty problem, which requires two pieces of information – acceptable risk and quantified uncertainty. The role of statistics is to quantify the uncertainty associated with the decision process. Statistical issues related to the setting of a criterion are different from issues related to compliance assessment. The presentation will focus on four topics, each with a key question and how to address the question.

1. The statistical interpretation of a numeric criterion – what are ideal and realizable criteria?

   The answer to this question will decide the interpretation of a criterion. For example, should the criterion be compare to the 90th percentile or the mean of the pollutant concentration distribution? An ideal criterion is defined for a population distribution parameter that cannot be measured. A realizable criterion is a statistic that can be calculated from data. The questions of averaging period and return frequency are about a realizable criterion.

2. Quantifying and reducing the uncertainty in a dose-response model – why we almost always expect a high level of uncertainty in a typical dose-response model and how to reduce the uncertainty?

   A typical dose-response model is developed based on data of limited sample size. These models often have high $R^2$ values but are inherently unreliable.

3. Conceptual confusion in the current reference condition guidelines – what is the reference distribution derived using current method?

   The current approach for deriving reference condition confuses the between and within site variances. The confusion made the resulting nutrient criteria meaningless.

4. Hypothesis testing for compliance assessment – why we should abandon statistical significance testing altogether?

Since the 1990s, not only we have accumulated a large sum of water quality monitoring data, we have also developed statistical methods that can better address the four questions raised in this presentation. We will discuss the use of the Bayesian hierarchical modeling approach as a unifying framework for these issues.

**#28**

**Name of Presenter:** [redacted]

**Organization:** [redacted]

**Address:** [redacted]

**Telephone:** [redacted]

**Email Address**: [redacted]

**Presenter Biography**:

[redacted]

**Presenter Title:** The Application of Continuous Monitoring to Chlorophyll-*a* Criteria Derivation, Evaluation, and Implementation

**Abstract:**

The need for scientifically defensible chlorophyll criteria has drawn increasingly more attention as jurisdictions tackle the challenge of controlling eutrophication, particularly in coastal estuaries. Algae blooms arising from nutrient pollution can harm aquatic life by altering the physicochemical properties of water. Elevated pH, low and excessively high dissolved oxygen levels, and reduced water clarity are problems commonly associated with nuisance algae growth, and thus make ideal endpoints for quantifying eutrophication. However, these parameters are highly variable and may not always be properly characterized through traditional monitoring. Using data collected over three years in three major tributaries of the Lower Chesapeake Bay, this presentation will illustrate the precise level of site characterization that continuous monitoring enables in the context of algae and physicochemical endpoints. The importance of site-specific nutrient-related criteria and sensitive monitoring techniques will be underscored by demonstrating the wide range of physicochemical exceedences observed among sites that are generally assumed to be impacted by comparable levels of nutrient overenrichmment. Additionally, an empirical model will be presented that uses high-frequency data to evaluate the protectiveness of existing chlorophyll-*a* criteria in relation to pH and DO percent saturation. A similar approach could be employed elsewhere to derive new criteria that are site-specific and based on best science.

**#29**

**Presenter Information**

**Presenter:** [redacted]

**Organization:** [redacted]

[redacted]

[redacted]

**Telephone:** [redacted]

**Email**: [redacted]

**Biographical Summary:** [redacted]

**Abstract**

**Title:** Ambient Water Quality Criteria are a Regulatory Tool, not a Risk Assessment Tool

[redacted]

Risk assessment combines information on site-specific environmental conditions, chemical exposure and toxicity to estimate the probability that site conditions will result in adverse effects on individuals, populations and communities. Ambient water quality criteria for the protection of aquatic life (aquatic life criteria) are not suitable for use in assessment of ecological risk because they are designed to be broadly applicable across the United States, and do not recognize or address site-specific conditions. Critical considerations in site-specific aquatic ecological risk assessment include characteristics of the water and water body (dissolved organic carbon, suspended organic material, hydrological dynamics), classification of the water body, the species present, species interactions, and the presence and role of other environmental stressors. This paper reviews data and case studies comparing conditions predicted by aquatic life AWQC to actual ecological conditions; these case studies demonstrate that aquatic life criteria do not effectively predict risk and are not appropriate for use in defining natural resource injury. Case studies suggest that provision for the derivation of site-specific aquatic life criteria could provide a practical means of improving water quality compliance. While the current and future aquatic life criteria may be appropriate as regulatory limits to meet the antidegradation policy of the Clean Water Act and may have utility in screening risk assessments, they are inadequate for prediction of ecological risk and resource injury, which require greater sensitivity to the multiple environmental conditions that mitigate or exacerbate adverse effects.

**#30**

Presenter: [redacted]

[redacted]

Biography: [redacted]

Topic: Approaches for adjusting water quality criteria based on physicochemical factors

Title: Developing water quality criteria that consider site-specific chemical factors using the Biotic Ligand Model

Co-Authors: [redacted]

Abstract: Among the recommendations from a Pellston workshop on a reevaluation of the state of the science for water quality criteria development[1] was that there should be greater use of mechanistic models for the prediction of bioavailability and chemical effects among different sites as a function of water-quality parameters. The biotic ligand model (BLM) has been developed as a generic modeling framework that aims to provide this mechanistic linkage between site-water chemistry and bioavailability and toxicity of metals to aquatic organisms. Metals typically have a complex chemistry that can be altered chemical factors such as pH, hardness cations, or the presence of natural organic matter. The BLM provides a mechanistic framework for considering how these chemical factors affect metal bioavailability in criteria development. In 2007 the U.S. Environmental Protection Agency (EPA) revised the ambient water quality criteria for copper using the BLM. Prior to this revision the copper criteria were based only on water hardness. However other water quality factors such as pH and the presence of natural organic matter can have pronounced effects on copper toxicity. Since hardness-based criteria do not consider these factors they may be under-protective in some conditions (e.g. low pH), and overprotective in other conditions (e.g. waters with elevated concentrations of natural organic matter). The BLM uses mechanistic relationships to determine how chemical factors can affect the bioavailability and toxicity of metals, and thereby provides a more accurate methodology for determining metals criteria over a wide range of conditions compared with hardness-based approaches. The BLM has been applied to a number of metals including aluminum, cadmium, cobalt, copper, nickel, lead, silver, and zinc, and BLM-based water quality criteria approaches for these other metals are also feasible. This presentation will review factors that affect metal bioavailability, discuss how the BLM addresses these factors, and discuss how consideration of bioavailability leads to more accurate water quality criteria. Bioavailability factors for a number of different metals will be reviewed and the implications on criteria development will be discussed.

[1] M.C. Reiley, W.A. Stubblefield, W.J. Adams, D.M. Di Toro, P.V. Hodson, R.J. Erickson, F.J. Keating Jr. 2003. Reevaluation of the state of the science for water quality criteria. SETAC Press, Pensacola, FL.

**#34**

Name of presenter: [redacted]
Organisation: [redacted]
Address, City, State, Zip: [redacted]
Telephone: [redacted]
Email address: [redacted]

**Changes to the Method for Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants**

[redacted]

Australia and New Zealand are nearing completion of a revision of their Water and Sediment Quality Guidelines. As part of this, there have been significant changes to the methods for deriving guideline values (GVs) for toxicants that have been adopted. These have already been applied to some 15 priority toxicants, with a further 30 currently being developed or revised.

Guideline value derivations use species sensitivity distributions (SSDs) of chronic toxicity data using the Burr Type III statistical distribution for ≥8 data points (updated software). Log-logistic fits are used for <8 points.

The acceptability of toxicity endpoints follows the order NEC, EC/IC/LCx where x≤10, BEC10, EC/LC15-20, NOEC, and NOEC estimated from MATC, LOEC or LC50 values. The use of NOECs is discouraged in favour of ECx values. Inclusion of non-traditional endpoints (e.g. behavioural) is allowed only if ecological relevance can be demonstrated.

Definitions of acute and chronic toxicity have been revised and the classification of toxicity tests altered. Macroalgal early life stage tests (ELS), e.g. fertilisation, germination and cell division, are now considered chronic; invertebrate early life stage tests, e.g. fertilisation and larval development are now considered chronic; and ≤7-d juvenile or adult fish and amphibian mortality tests are now considered acute.

When using SSDs to derive GVs, the order of preference for toxicity data is for chronic data for ≥8 species (although there is an aspirational target of ≥15 species); chronic + converted acute data for ≥8 species; chronic data for 5-7 species; chronic + converted acute data for 5-7 species; converted acute data for ≥8 species; and converted acute data for 5-7 species. Data must comprise at least 4 taxonomic groups.

In toxicity testing, although some indication of repeatability is desirable, it is recommended that, rather than replication of concentrations, more concentrations should be obtained at the lower end of the concentration-response relationship (i.e. below a 50% effect) to reduce uncertainties and improve the precision of statistical estimates of toxicity to be used in the SSD. This is consistent with the recommended move away from hypothesis testing to derive NOECs, to the derivation of NEC and ECx values.

A new reliability classification for SSD-derived GVs has been developed based on (i) the hierarchy of acceptable data, (ii) the sample size, and (iii) a visual estimation of goodness of fit. Very high reliability GVs have a good fit for ≥15 chronic data points. A good fit and ≥8 data points is classified as high reliability. The use of assessment factors is a last resort, with sample size considered inadequate and reliability unknown.

The concept of using GV exceedance as a trigger for further investigation is retained, however, a weight-of-evidence approach is being adopted as part of the overall water quality assessment framework. This allows for the inclusion of ecology, bioaccumulation, biomarkers and other lines of evidences in addition to chemistry and toxicity, to provide greater certainty in assessment decisions.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted]

**#35**

[redacted]

[redacted] bio:

[redacted]

**Challenges to improve the utilization and applicability of water quality criteria for metals**

Metals and inorganic chemicals present particular challenges for setting environmental quality standards, given that many of the regulatory frameworks around the world were designed to evaluate synthetic organic chemicals. The US EPA Framework for Metals Risk Assessment (MF) and the ICMM's Metals Environmental Risk Assessment Guidance (MERAG) lay out unique physical/chemical characteristics and toxicological considerations of metal behavior in aquatic systems that need to be considered in developing WQC. The EPA Office of Water (OW) has been a leader in the application of metal-specific tools, such as the Biotic Ligand Model (BLM), in the regulation of metals. This paper will evaluate the MF and MERAG for commonalities in addressing metals risk assessment issues and EU advancements in the regulation of metals through the Water Quality Framework Directive. EPA OW may be able to leverage from these documents to incorporate additional metal-unique tools for metal WQC.

Current research, to include tiered approaches, simplified BLM calculations, and mixtures, will be reviewed in the context of a path forward to increase the utilization of these methods in the regulation of metals in aquatic systems.

**#36**

**“Making Standards Stick – Derivation and Implementation of Water Quality Standards”**

## ABSTRACT

Environmental thresholds like Environmental Quality Standards (EQSs) or Water Quality Criteria (WQC) play a key role in assessing the risks from chemical pressures in the environment. These thresholds are developed through regulatory frameworks like the Water Framework Directive and Marine Strategy Framework Directive in Europe and National Ambient Water Quality Criteria in the US.

Whilst there have been numerous developments in the underpinning science and the methods used to derive these thresholds in recent years, such advances are insufficient to deliver actual environmental benefit unless they are coupled with practical solutions to the challenges of implementation. Confidence about the actions we take (or don’t take) in response to decisions about chemical status is also important to the general public as the financial burdens associated with programmes to fund environmental improvements increasingly fall on the taxpayer. My presentation focuses on the challenges of implementing environmental standards so they will achieve their intended goals in an affordable way.

Bridging the gap between research and regulatory implementation to deliver environmental improvement is not always straight-forward. Often the focus of attention is the derivation of the numerical value of the standard (the concentration of a chemical in an environmental matrix like water or sediment). Particular implementation difficulties include dealing with uncertainty, the setting of quality standards beyond the sensitivity of existing analytical methods, implementing standards for metals whose toxicity is affected by local conditions, dealing with backgrounds, and the introduction of standards for highly lipophilic substances that biomagnify.

Field validation of standards which are derived predominantly from laboratory testing is rare, but can be very useful in corroborating our predictions and demonstrating the validity of proposed standards to stakeholders. Such lines of evidence have an important part to play in setting environmental standards, alongside laboratory- based effects data.

There are also important issues about the statistical expression of standards, and the way decisions about compliance are reached, and these are just as important as the numerical value of the standard. Problems of implementation are often overlooked by researchers and regulators since it has not prevented the setting of

the standard, and the scientific and practical questions around implementation of a standard can be underestimated, or missed completely.

In my presentation I will draw on experience gained in Europe, particularly through the Water Framework Directive, but the principles apply to the derivation and implementation of chemical thresholds irrespective of the regulatory regime that led to their derivation. My presentation will suggest a greater focus on the implementation of standards to help ensure a proper balance between the needs of the environment and the technical and financial burden on regulators and stakeholders.

[redacted]

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

**B I O G R A P H Y**

[redacted]

[redacted]

[redacted]

[redacted]

Biography:

[redacted]

[redacted]

[redacted]

#37

**Name of Presenter:** [redacted]

[redacted] **Organization:**

[redacted]

**Address:** [redacted]

**Telephone:** [redacted]

**Email Address:** [redacted]

**Presenter Biography:** [redacted]

**Presentation Title**: Using Weight of Evidence to Evaluate Potential Effects of Trace Organics on Aquatic Life

**Abstract**: Thousands of trace organic compounds (TOrCs) are regularly released into the aquatic environment, and little is known of the effects of this complex mixture, although targeted studies suggest exposures could have important impacts on aquatic communities. Because of uncertainty concerning potential risks, management actions might focus on the wrong sites and/or compounds, and might not yield expected improvements. Establishing a causal linkage between wastewater treatment plant (WWTP)-related TOrC exposures and biological impairments is challenging due to the unsettled state of science with respect to which TOrCs are endocrine disruptors; what fraction of the TOrCs may be bioavailable; how TOrCs interact with other

stressors; and whether the presence of TOrCs is a reliable indicator of aquatic life impairment. All these confounding considerations suggest there may be very different exposure and effects in an actual receiving waterbody than is indicated by screening models and laboratory studies. The approach we are piloting for the Water Environment Research Foundation screens WWTP discharges and assigns each to a risk-and action-based Tier of Concern, based on documented impairments and site stressors, level of wastewater treatment, dilution by receiving water, concentrations of TOrC indicators, and site-scale, screening-level evaluations of potential risk to aquatic life. For sites where existing data suggest moderate or higher levels of concern, field work is recommended to support a weight of evidence evaluation, including discharge and receiving water sampling, fish and macroinvertebrate bioassays, toxicity identification evaluations, and habitat assessments. The goal of the weight of evidence assessment is to allow WWTPs to better assess risk and potential causation to inform monitoring planning and wastewater treatment investments. The presentation will report on pilot field studies conducted at three sites in summer 2014, and explore implications for Aquatic Life Criteria. This project is part of an ongoing long-term research commitment by the Water Environment Research Foundation to  better understand and manage impacts of Trace Organics on aquatic ecosystems.

#39

[redacted]

Presenter Biography

[redacted]

**Topics Addressed by the Presentation and Areas I can contribute to:**

- Use of statistical tools for developing criteria
- Incorporation of field or meso/microcosm data to derive or evaluate criteria
- Ecological representation of data used to develop criteria
- Selecting and prioritizing pollutants of concern
- Incorporating weight-of-evidence approaches
- Use of alternative biological endpoints in criteria development
- Approaches for adjusting criteria based on physicochemical factors

## The Interaction of Hydrology, Habitat, and Biota in Warmwater Urban Streams – Implications for Water Quality Criteria Development

The term "Urban Stream Syndrome" has been used to characterize streams exhibiting the symptoms of urbanization. Within Texas streams in urban areas have been subjected to various stressors including altered hydrology, channel modification, reduction in instream habitat, loss of riparian cover, and impaired water quality. Many of these alterations have resulted in creating conditions that facilitate the invasion of exotic species that further affect the distribution and abundance of native taxas. The interaction of these various factors on water quality often result in multiplicative impacts not predicted by numerical water quality criteria alone and not easiliy reversed.

Data from various stream surveys conducted over the last 30 years within southeast Texas are used to describe temporal and spatial trends in stream fish communities associated with increasing urbanization and associated factors. We have observed an overall increase in the number and density of invasive exotic aquatic species in tidal and coastal freshwater streams experiencing urbanization. Sources of exotic species include aquarium releases, intentional stocking, and escapes from aquaculture facilities. During this period of time there has been a concurrent conversion of normal varying seasonal stream flows to constant elevated effluent dominated base flows, elevated storm flows, and reduced instream habitat.

The altered hydrology and water quality has favored the establishment of invasive exotic species including Applesnails, Tilapia and Suckermouth catfish. These species share many life history attributes including high fecundity, omnivory, and tolerance of impaired water quality. These life history traits allow these introduced species to more efficiently exploit these altered stressed environments and out-compete more sensitive native fish species under water quality conditions that may marginally comply with numerical water quality criteria. The result is an amplified negative affect on native aquatic organisms over predicted impacts from water quality impairment alone. This regime shift from healthy native aquatic communities to altered communities dominated by a few tolerant exotic and native species occurs at a "system degradation threshold (SDT)" not predicted by numerical water quality criteria alone. Furthermore reversal of this sequence back to the original state cannot be achieve by a simple improvement by one of two variables including water quality.

Attempts to protect and restore native aquatic communities in these urban streams will require a comprehensive watershed approach that includes reestablishment of traits associated with natural stream morphology and instream habitat, reestablishment of riparian vegetation and reduction in pollutant loads to attain concentrations potentially below existing numerical water quality criteria. However in some warmwater tidally influenced and coastal freshwater streams there is evidence that native populations of biota may have evolved to survive in naturally occuring adverse water quality conditions (e.g. hypoxia). Restoration of urban and rural warmwater streams will require careful evaluation of current and proposed numeric criteria, watershed scale activities and site specific actions to be successful. Currently comprehensive integrated policies, tools and institutions are lacking that integrate the various scales of activities necessary to accomplish stream restoration that includes reestablishment of native aquatic communities as a goal.

#40

**Assessment of Key Aspects and Deficiencies of the 1985 Guidelines in Directing Pollution Reduction Resources for Aquatic Life Protection**

[redacted]

The development and expression of aquatic life protection criteria serves as the foundation for the water quality based permitting program implemented throughout the United States. Consequently, literally a trillion dollars in public and private environmental resources expenditures is dependent on the accuracy and reasonableness of the criteria development program. The 1985 *Guidelines fr Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses* (hereafter " 1985 *Guidelines*") has played a pivotal role for ensuring that scientifically defensible criteria are generated based on a sufficient technical database. Several aspects of the 1985 *Guidelines* are still essential for ensuring proper program implementation, while others have clearly become dated and in need of revision.

The author has over 35 year's experience in water quality based program implementation and numeric/narrative criteria development in over 45 states. This presentation would cover key aspects of the 1985 *Guidelines* that have been instrumental in ensuring pollution reduction requirements are properly targeted under the NPDES program by federal, state and local governments. It would also highlight areas where existing procedures are deficient or are still encountering "integration" issues with delegated state program implementation, thereby potentially misdirecting ecological resource expenditures and undercutting attainment of Clean Water Act objectives. The topics to be covered will include:

- Review of essential provisions of the 1985 *Guidelines*
- Identification of outdated/misdirected 1985 *Guideline* provisions
- Assessment of Margin of Safety – over and under regulation
- Approaches to aquatic life protection under uncertainty
- Suggested *Guideline* components and clarifications for ensuring consistent NPDES/TMDL program implementation

[redacted]

[redacted]

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[redacted]